Mechanism design for federated sponsored search auctions

Citation for published version:
Ceppi, S, Gatti, N & Gerding, EH 2011, Mechanism design for federated sponsored search auctions. in *25th AAAI Conference on Artificial Intelligence (AAAI).* AAAI Press, pp. 608-613.
<http://www.aaai.org/ocs/index.php/AAAI/AAAI11/paper/view/3576>

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
25th AAAI Conference on Artificial Intelligence (AAAI)

An Automated Mechanism Design Approach for Sponsored Search Auctions with Federated Search Engines

S. Ceppi and N. Gatti

Dipartimento di Elettronica e Informazione, Politecnico di Milano
Piazza Leonardo da Vinci 32, I-20133 Milano, Italy
{ceppi,ngatti}@elet.polimi.it

Abstract. Advertising mechanisms for search engines (i.e., sponsored search auctions) have recently received a lot of attention in the scientific community. Advertisers bid on keywords and, when a user enters keywords for her search, the search engines uses an auction mechanism to select the list of sponsored links to display alongside the search results. In this paper, we make a first attempt to extend the currently available mechanisms for sponsored search auctions to the new paradigms of search computing. According to them, multiple federated domain-specific search engines are integrated by a special search engine (called integrator). The user can enter a multi-domain query that is decomposed by the integrator in single-domain queries and these are singularly addressed to the most appropriate domain-specific search engine. The integrator merges the search results. We propose a business model for this scenario and we develop an economic mechanism for it resorting to the automated mechanism design approach.

1 Introduction

Sponsored search auctions [1, 2] play a prominent role in Internet advertising, generating more than 90% of the search engines' revenues. A large number of theoretical/practical works can be found in the very recent literature. Nevertheless, this market is still largely unexplored and a number of problems are currently open. The functioning of sponsored search auctions is simple. When a user enters keywords into a search engine, several sponsored links related to the entered keywords are displayed alongside the search results, e.g., see [3]. The search engine chooses the sponsored links to display and the ranking over them by using an auction mechanism where the bidders are advertisers and the item over which they bid are keywords. The payment scheme is the *pay-per-click*, i.e., the advertiser pays the search engine only after a user has clicked on its sponsored link.

The most employed auction mechanism in sponsored search auctions is the *generalized second price* (from here on GSP) [4] that is an *ad hoc* extension of the Vickrey auction (from here on VA) [2] to the setting where a set of ranked objects is being sold. In the VA, a winner pays the second highest bid, while, in the GSP, each winner pays an amount equal to its next highest bid. However, as shown in [4], in the GSP the truth telling strategy is not (generally) optimal for the players, as instead it is in the VA. The exact generalization of the VA to the above settings, satisfying the property that the truth-telling strategy is optimal, is similar to the GSP except for the definition of the payments. Although this last mechanism is strategy-proof and could assure a higher degree of the outcome stability, it is not currently adopted in real-world applications.

The available economic mechanisms for sponsored search auctions effectively work with the major general-purpose search engines. However, the recent advancements in the search computing field lead to the definition of novel searching paradigms that rise new challenges and that require extensions of the available auction mechanisms [5, 6]. The main general-purpose search engines crawl the Web and index Web pages, finding the best pages for each specific list of keywords with excellent precision. Anyway, the so-called "deep Web" contains information that is more valuable than that contained in single Web pages and the current general-purpose search engines are not able to discover it. The development of new searching paradigms able to address more complex searches than those addressed to the current search engines and to discover deeper information is currently one of the most interesting challenges in the search computing field. In particular, the emerging paradigm is based on the integration of heterogeneous data sources. According to that, a special search engine (from here on *integrator*) integrates the results produced by multiple domain-specific search engines, e.g., see [7]. The basic idea is the following. The user's search is a multi-domain query. Each multi-domain query is automatically decomposed by the integrator in multiple single-domain queries and each of them is addressed to the most appropriate domain-specific search engine. Obviously, when a query addresses a specific domain, domain-specific search engines works better than general-purpose ones. Once the integrator has received the search results from all the domain-specific search engines, it aggregates them in a unique result. This is shown by using *ad hoc* interfaces that allow the user to explore the search results, adding/removing domains and thus refining the search itself, e.g., see [9].

The search computing field is working exclusively on the searching techniques and is neglecting the business model behind the above scenario (e.g., what kind of contracts will be drawn up between the integrator and the domain-specific search engines?). Currently the commercial use of the search results produced by a search engine is ruled by a contract between the search engine and the publisher prescribing that the publisher must display the list of sponsored links produced by the search engine, e.g., as in [10]. Once a user clicks on a sponsored link, the search engine receives the payment from the corresponding advertiser and gives part of it to the publisher. The payment ratio kept by the search engine is defined by a commercial contract and it is independent of the specific search. On the one hand, the basic idea behind this business model can be "naturally" applied to the above scenario. On the other hand, the contracts between the publisher (in our case the integrator) and the search engines (in our case the domain-specific search engines) must be reconsidered keeping into account that each search engine plays a role in the search process. Our opinion is that the contracts between the integrator and the search engines should be drawn up dynamically, depending on the specific search and, in particular, on the contribution provided by the specific search engine to the search.

In this paper we propose an economic mechanism [11] to rule the contracts between the integrator and the domain-specific search engines. In our proposal, the integrator receives the lists of sponsored links from the domain-specific search engines and merges them in a unique list. In the merging process, the integrator keeps into account the advertisers' bids and click probabilities related to the list of each domain-specific search engine in order to generate the list of sponsored links that gives the largest expected utility. Being the information on the advertisers' bids and click probabilities private for each domain-specific search engine, we must produce the appropriate incentives to the domain-specific search engines not to misreport such a information. We formulate this problem as a single-stage mechanism design problem [2] and we discuss the desired properties. We show that the domain-specific search engines present interdependent valuations due to the aggregation of their information. We study it by using the automated mechanism design approach [12]. It provides a flexible tool to design mechanisms on-the-fly and allows one to customize each problem by varying different objective functions and adding/removing possible constraints over the contracts. However, the hardness of solving an automated mechanism design problem allows us to solve in exact way only small settings with a few of search engines and advertisers. For large settings, approximate (anytime) algorithms can be developed to produce a sub-optimal solutions by a given deadline.

Finally, we remark that the possibility to integrate multiple lists of sponsored links provides, in our opinion, two advantages. First, the integrator can target at best the advertisement to the user by exploiting multiple information sources (i.e., the domain-specific search engines) and the user's feedback during her exploration of the search. Second, this paradigm allows domain-specific search engines to federate together and to be real competitors to the major general-purpose search engines. This could open new economic opportunities for online advertising.

The rest of the paper is structured as follows. In Section 2, we discuss the state of the art related to the sponsored search auctions, the multi-domain search computing, and the automated mechanism design. In Section 3, we propose a business model for the scenario we study, we formally state an economic mechanism, we discuss the desired properties, and we formulate the problem of designing the mechanism as an automated mechanism design problem. In Section 4, we discuss some examples. Section 5 concludes the paper.

2 State of the Art

We introduce a formal model of sponsored search auctions in Section 2.1, we discuss the multi-domain search computing paradigm in Section 2.2, and we survey the idea behind the automated mechanism design in Section 2.3.

2.1 Sponsored Search Auctions

The formal model of a sponsored search auction is constituted by m *items* (i.e., the ranked set of slots for sponsored links given a specific keyword) sold by the *auctioneer* (i.e., the search engine) and by a set of *bidders* (i.e., the advertisers) $N = \{1, 2, \dots, n\}$ where $n \geq m$. Each advertiser can submit a bid constituted by a value per click on the advertisement for a keyword. The bid is unique for all the slots. Being the payment scheme pay-per-click, each advertiser pays nothing if its sponsored link is displayed but not clicked by the user. Instead, in the case the user clicks on the link, the advertiser is required to pay an amount of money that is non-larger than its bid. The exact value of the payment is carried out by the auction mechanism.

The search engine assigns to each bidder a click probability called *click-through-rate* (from here on CTR) [2]. CTR depends on various factors including the probability that users click on advertisement, the relevance of the bidders' advertisement, and so on. Formally, we denote by $\alpha_{i,j}$ the probability that the advertiser i's sponsored link is clicked when it is displayed on the j-th highest slot. Usually, these probabilities are supposed to be separable into two independent components, where the first component refers only to the advertiser and the second component refers only to the position of the slot. Formally, $\alpha_{i,j} = \alpha_{a_i} \cdot \alpha_{r_j}$ where α_{a_i} is the probability that advertiser i is clicked independently of the specific slot in which its sponsored link is displayed and α_{r_j} is the probability assigned to the j-th highest slot independently of the specific advertiser. The common assumption is that $\alpha_{r_1} > \alpha_{r_2} > \ldots > \alpha_{r_k}$. Currently, Google ranks the advertisements by using a separable CTR.

The GSP auction extends the VA as follows. We call b_i the bid submitted by advertiser i. The auctioneer ranks the bids in decreasing order in the value $\alpha_{a_i} \cdot b_i$. For the sake of simplicity suppose that, given α_{a_i} and α_{a_h}, if $i < j$, then $\alpha_{a_i} \cdot b_i > \alpha_{a_h} \cdot b_h$. The advertiser with $\alpha_{a_1} \cdot b_1$ will be displayed at the first slot, the advertiser with $\alpha_{a_2} \cdot b_2$ will be displayed in the second slot, and so on. Once the sponsored link displayed in the i-th position is clicked, the i-th advertiser pays $p_i = \frac{\alpha_{a_{i+1}} \cdot b_{i+1}}{\alpha_{a_i}}$ to the search engine (while the other search engines' payments are zero).

2.2 Multi-Domain Search Computing

The new advancements produced in the field of the search computing are directed to fill the gap between general purpose search engines and domain-specific search engines. General purpose search engines work well in finding Web pages related to the entered keywords, but are unable to find information spanning multiple topics. Domain-specific search engines work well in finding structured information spread on multiple Web pages related to a specific domain, but their expertise is clearly limited to a given domain. An expert user can perform several independent searches and then manually combine the results, but the missing aspect is the ability of joining the results of each search process so as to build a collective answer.

The emerging search computing paradigm prescribes that a user can compose her search as a multi-domain query and this is addressed to a federated search platform. Each single-domain query is addressed to the most appropriate domain-specific search engine. The search results produced by the domain-specific search engines are merged by an integrator through syntactic and/or semantic joint methods [13]. The merging of the search results is only the first feature provided by the integrator. Indeed, this allows a user to explore the search by changing modularly the dimensions of the multi-domain query. In particular, a user can add or remove dimensions, can manipulate results via composition and aggregation, and can reorder results [14].

Search results are shown through tables where the columns are the dimensions and the rows are items of the search. In the cells of the tables, information related to the specific item and dimension is reported. An example is Google Squared [9]. The definition of a business model for federated search engines and the integrator is currently an open problem.

2.3 Automated Mechanism Design

Classical mechanism design provides general mechanisms, satisfying some notion of non-manipulability and maximizing some objective. The most famous general mechanisms, VCG and dAGVA [2], only maximize social welfare. For almost all the other social choice functions, there is no known mechanism that implements them. For example, revenue-maximizing mechanisms are only known for very restricted settings, such as the Myerson's expected revenue maximizing auction for selling a single item, and the Maskin and Riley's expected revenue maximizing auction for selling multiple identical units of an item [2]. In practice, a designer often has prior information over agents' types and only needs to design a mechanism suitable for her particular context. In the automated mechanism design approach, a mechanism is designed automatically for the specific preference aggregation problem. The mechanism design problem can be formulated as an optimization problem where the input is characterized by the number of agents, the agents' possible types (preferences), and the aggregator's prior probability distributions over the agents' types and the output is a non-manipulable mechanism that is optimal with respect to some objective.

3 The Economic Mechanism

We propose a business model in Section 3.1 for multi-domain search computing scenarios in which domain-specific search engines are federated, we formulate an economic mechanism supporting our business model in Section 3.2, we discuss the desired properties of our mechanism in Section 3.3, and we state the problem of designing the social choice function and the payments in our mechanism as an automated mechanism design problem in Section 3.4.

3.1 The Proposed Business Model

The commercial use of the search results is currently ruled by a contract between the search engine and the publisher prescribing that the publisher must display the list of sponsored links alongside the search results, e.g., see [3]. If the user clicks on a sponsored link, the corresponding advertiser pays the search engine which, in its turn, gives a part of the revenue to the publisher. The ratio of revenue kept by the search engine does not depend on the specific search and is established by the contract. The values of the advertisers' bids and the corresponding click probabilities constitute private information of the search engine and are hidden to the publisher.

This business model can be easily applied to the case in which an integrator merges the search results of multiple search engines. The idea is:

- the integrator merges the lists of sponsored links returned by each single domain-specific search engine,
- the payments from the domain-specific search engines to the integrator depends on the specific search.

The crucial issue is the development of techniques that allows the integrator to produce the list that maximizes a given objective function (e.g., the expected revenue of the integrator or the expected revenue of a combination of specific-domain search engines). We propose an economic mechanism to govern the merging of the lists where:

- the domain-specific search engines communicate to the integrator their private information (values of the bids and click probabilities) concerning the sponsored links related to their own list;
- the integrator produces an estimation of the click probability for each sponsored link as a function of the received click probabilities (e.g., averaging the click probabilities of a sponsored link over the different search engines);
- the integrator selects the list of sponsored links in order to maximize a given objective function (e.g., the integrator's expected utility or the cumulative expected utility) and produces the appropriate incentives (i.e., payments) for each domain-specific search engine to make them not misreport their true values;
- the integrator keeps into account how the user explores the search results in estimating the click probabilities and produces a new list of sponsored links every time the user add/remove domains.

In what follows, we formally state the economic mechanism.

3.2 The Formal Mechanism

We consider a direct mechanism [2] $\mathcal{M}(X, S, \Theta, V, f, p)$ where the agents are the domain-specific search engines (from here on we omit "domain-specific") and the integrator acts as auctioneer. We denote by X the set of alternatives. For the sake of presentation, we formally define X below, after having defined S and

Θ. We denote by S the set of search engines. We denote by A the overall set of advertisers and by $A(s)$ with $s \in S$ the set of advertisers appearing in the list of sponsored links of search engine s. Each advertiser $a \in A(s)$ is characterized by a bid b and a click probability α that are private information for search engine s. The type $\theta_s \in \Theta_s$ of search engine s specifies a value for b and a value for α for all the advertisers $a \in A(s)$. Set Θ is composed by all the sets Θ_ss and θ denotes the profile of search engines' types. We assume that we have a probabilistic prior over Θ_s and we can represent it as a set of independent probability distributions, each over a specific advertiser $a \in A(s)$. In particular, we denote by $\Theta_{s,a}$ the set of possible types of advertiser a appearing in the list of sponsored links of search engine s and we denote by $\theta_{s,a} \in \Theta_{s,a}$ the type. We introduce the functions $b(\theta_{s,a}) : \Theta_{s,a} \rightarrow \mathbb{R}$, returning the bid submitted by advertiser a to search engine s related to type $\theta_{s,a}$, and $\alpha(\theta_{s,a}) : \Theta_{s,a} \rightarrow [0, 1]$, returning the click probability of advertiser a in the sponsored link list of search engine s related to type $\theta_{s,a}$. Generally, an advertiser a can appear in the lists of more than one search engine with different values of bid and click probability, i.e., $b(\theta_{s,a})$ can be different from $b(\theta_{s',a})$, as well as $\alpha(\theta_{s,a})$ can be different from $\alpha(\theta_{s',a})$. We denote by $\omega(\theta_{s,a})$ the probability that the actual type of advertiser a for search engine s is $\theta_{s,a}$. Therefore, the type of search engine s is a tuple specifying the type related to each advertiser $a \in A(s)$, e.g., $\theta_s = (\theta_{s,1}, \ldots, \theta_{s,|A(s)|})$. The probability $\omega(\theta_s)$ related to θ_s is defined as $\omega(\theta_s) = \prod_{\theta_{s,a} \in \theta_s} \omega(\theta_{s,a})$.

Now we focus on set X. An alternative $x \in X$ specifies a winner for each slot of the list of sponsored links displayed by the integrator. We assume that the number of available slot is fix and it is equal to k. A winner is identified by a pair (s, a), that is advertiser a related to the sponsored link list of search engine s. This is because the same advertiser $a \in A$ can appear in the sponsored link lists of different search engines. We need to specify the search engine to which the sponsored link belongs because such a search engine will be paid by the advertiser and a may have submitted different bids to different search engines. Formally, $x = \langle (s, a), \ldots, (s', a') \rangle$, where the first element of x specifies the winner of the first slot, the second element of x specifies the winner of the second slot, and so on. The unique constraint is that a sponsored link can appear only in one slot, that is, for all a and a' appearing in x in different positions, we have $a \neq a'$.

We denote by V the set $V = \{v_s : s \in S\}$ where $v_s : X \rightarrow \mathbb{R}$ denotes the valuation function of search engine s. Given x, if s does not appear in x, then $v_s(x) = 0$. Instead, if s appears in x, $v_s(x)$ returns the s's expected valuation over x defined as: for each $(s, a) \in x$, the s's expected valuation is the product between the a's click probability and the valuation that s receives when the a's sponsored link is clicked. Before formally stating $v(x)$, we focus on these two elements. First, we consider the a's click probability. It is a function of $\alpha(\theta_{s',a})$ for all s' where $\theta_{s',a}$s are the reported types. That is, the integrator produces an estimation of such click probability, denoted by $\overline{\alpha}(a)$, aggregating the click probabilities over the advertiser a of all the search engines s such that $a \in A(s)$. In estimating $\overline{\alpha}$, the integrator can exploit several parameters, e.g., it can assign different weights to different search engines or excluding search engines, once the user has removed the corresponding dimension. In our work we use a simple estimator: the average $\overline{\alpha}(a) = \frac{\sum_{s \in S: a \in A(s)} \alpha(\theta_{s,a})}{|A|}$. Second, we consider the valuation that s receives when the a's sponsored link is clicked. In this work we assume that advertiser a pays to s exactly its bid. Essentially, we assume a first-price approach. We make this assumption for simplicity because we are focusing only on the interaction between the integrator and the search engines. In future works, we shall consider also the interaction between the search engines and the advertisers. Now, we are in the position to formally state $v_s(x)$ as $v_s(x) = \sum_{(s,a) \in x} \overline{\alpha}(a) \cdot b(\theta_{s,a})$ where $\theta_{s,a}$s are the true types of s.

In mechanism $\mathcal{M}$, f and p define respectively the social choice function and the search engines' payments. More precisely, f is a function $f : \Theta \to X$ that given the type of all the search engines returns an alternative, while $p : \Theta \to \mathbb{R}^{k \cdot |S|}$, where k is the number of slots, returns the payment for each search engine for each situation in which one sponsored link is clicked. We use a quasi-linear setting where the utility of a search engine is equal to the difference between its valuation and the payment. We want to design f and p such that $\mathcal{M}$ satisfies a set of properties.

3.3 Required Properties

Before discussing the properties we require that our mechanism satisfies, we underline that, in the general case, search engines in mechanism $\mathcal{M}$ present interdependent valuations [15, 16]. This is because $v_s(x)$ depends on $\overline{\alpha}(a)$ that, in its turn, depends on $\theta_{s',a}$s for all s'. Exclusively when $A(s) \cap A(s') = \emptyset$ for all $s \neq s'$, $\overline{\alpha}(a)$ depends only on the type of the search engine s such that $a \in A(s)$ and therefore the search engines' valuations are not interdependent. We require the following properties.

(Ex-post) Individual rationality. For every x such that $f(\theta) = x$, we require that, for every realization of x, the utility of all the search engines is non-negative. This requires that, given x, whenever a sponsored link a related to search engine s is clicked, s does not pay the mechanism more than $b(\theta_{s,a})$, while the payments of all the other search engines s's are non-positive.

(Ex-post Nash and Bayesian) Incentive compatibility. We require the implementation of f either in *ex-post* Nash or in Bayes-Nash equilibrium. Therefore, we require that each search engine reporting its true type is an optimal strategy. (We use *ex-post* Nash implementation instead of dominant strategy implementation because in our problem valuations are interdependent. We recall that *ex-post* Nash and dominant strategy implementations are always the same except when the valuations are interdependent.)

(Ex-post) Weak budget balance. For every x such that $f(\theta) = x$, we require that, for every realization of x, the cumulative payments of the search engines is non-negative. This requires that, given x, whenever a sponsored link a appearing in the sponsored link list of search engine s is clicked, the sum of search engines' payments excluded s is not smaller than $-b(\theta_{s,a})$. The revenue of the integrator is equal to the cumulative payments of the search engines.

Optimality. We consider several objective functions: the maximization of the *ex-ante* expected utility of the integrator (defined as the sum of the expected payments of the search engines), the maximization of the *ex-ante* cumulative expected valuations of the search engines, the maximization of the *ex-ante* cumulative expected utilities of the search engines, and the maximization of the *ex-ante* expected utility of a specific search engine. The choice of the objective function depends on the specific contract.

We remark that in the general case (interdependent valuations) our mechanism cannot be efficient. Indeed, with interdependent valuations and multiple signals a one-stage mechanism may not be incentive compatible and efficient [17] (with two-step mechanism is instead possible to have efficient incentive compatible mechanisms [16]; we shall explore this option in future works). It can be easily shown that even in the basic case in which there are two search engines and the same advertiser for both search engines, and only the click probabilities are uncertain, a one-stage mechanism may not be incentive compatible and efficient. Indeed, in the case each player has a single signal, three conditions need to be satisfied in order to make a mechanism incentive compatible and efficient, see [17]. One of these requires that $\frac{\partial v_s}{\partial \theta_s} > \frac{\partial v_{s'}}{\partial \theta_s}$ for all s, s'. Generally this condition is not satisfied in our basic case. As a result, f cannot be easily defined as the argument maximizing the social welfare, as instead it is possible for efficient mechanism.

3.4 The Automated Mechanism Design Formulation

We formulate our mechanism as an automated mechanism design [12] problem. We represent f as a collection $f_{s,a,\theta,r} \in \{0,1\}$ where $f_{s,a,\theta,r} = 1$ means that advertiser a related to search engine s is assigned position r when type profile of the search engines is θ. Index r belongs to the range $R = \{1,\dots,k\}$. For simplicity, for $a \in A \setminus A(s)$ we set $f_{s,a,\theta,r} = 0$, $b(\theta_{s,a}) = 0$, and $\omega(\theta_{s,a}) = 0$. We introduce the constraints in a mathematical programming fashion. Initially, we constrain every sponsored link a to appear at most in one position r:

$$\sum_{r\in R}\sum_{s\in S: a\in A(s)} f_{s,a,\theta,r} \leq 1 \qquad \forall \theta \in \Theta, \forall a \in A \tag{1}$$

We constrain that for each position r there is exactly one sponsored link:

$$\sum_{s\in S}\sum_{a\in A} f_{s,a,\theta,r} = 1 \qquad \forall \theta \in \Theta, \forall r \in R \tag{2}$$

We denote by $p_{s,\theta,r}$ the payment of the search engine s when the r-th sponsored link is clicked and the type profile is θ. To make the mathematical programming formulation easier, we divide $p_{s,\theta,r}$ in payments concerning the single advertisers, one for each of them. We denote these payments by $p_{s,a,\theta,r}$ and we define $p_{s,\theta,r} = \sum_{a\in A(s)} p_{s,a,\theta,r}$ for all $s \in S, \theta \in \Theta$, and $r \in R$. The *ex-post* individual rationality constraints make each search engine to pay no more than its valuation, formally, we have:

$$p_{s,a,\theta,r} \leq b(\theta_{s,a}) \cdot f_{s,a,\theta,r} \qquad \begin{array}{r}\forall s \in S, \forall a \in A(s),\\ \forall \theta \in \Theta, \forall r \in R\end{array} \tag{3}$$

We require further that, if the sponsored link of advertiser a is displayed at the r-th position when the type profile is θ, then its payment is non-negative, formally, we have:

$$p_{s,a,\theta,r} \geq -M \cdot \left(1 - \sum_{a' \in A(s)} f_{s,a',\theta,r}\right) \qquad \begin{array}{r} \forall s \in S, \forall a \in A(s), \\ \forall \theta \in \Theta, \forall r \in R \end{array} \tag{4}$$

where M is an arbitrarily large number. With abuse of notation we denote by $\overline{\alpha}(a,r)$ the probability that the integrator assigns to a when it is displayed at the r-th position. We represent $\theta = (\theta_s, \theta_{-s})$ where θ_s is the type profile of search engine s and θ_{-s} is the type profile θ once excluded θ_s. The *ex-post* Nash incentive compatibility constraints are:

$$\begin{array}{l} \sum_{r\in R}\sum_{a\in A}\Big(b(\theta_{s,a}) \cdot f_{s,a,(\theta_s,\theta_{-s}),r} - p_{s,a,(\theta_s,\theta_{-s}),r}\Big) \cdot \bar{\alpha}(a,r) \geq \\ \quad \sum_{r\in R}\sum_{a\in A}\Big(b(\theta_{s,a}) \cdot f_{s,a,(\theta'_s,\theta_{-s}),r} - p_{s,a,(\theta'_s,\theta_{-s}),r}\Big) \cdot \bar{\alpha}(a,r) \end{array} \qquad \begin{array}{r} \forall s \in S, \\ \forall \theta \in \Theta, \\ \forall \theta'_s \in \Theta_s \end{array} \tag{5}$$

The Bayesian incentive compatibility constraints are:

$$\begin{array}{l} \sum_{\theta_{-s}}\sum_{r\in R}\sum_{a\in A}\bigg(\Big(b(\theta_{s,a}) \cdot f_{s,a,(\theta_s,\theta_{-s}),r} - \\ \qquad\qquad - p_{s,a,(\theta_s,\theta_{-s}),r}) \cdot \bar{\alpha}(a,r)\Big) \cdot \prod_{s'\in S/\{s\}} \omega(\theta_{s'})\bigg) \geq \\ \sum_{\theta_{-s}}\sum_{r\in R}\sum_{a\in A}\bigg(\Big(b(\theta_{s,a}) \cdot f_{s,a,(\theta'_s,\theta_{-s}),r} - \\ \qquad\qquad - p_{s,a,(\theta'_s,\theta_{-s}),r}\Big) \cdot \bar{\alpha}(a,r)) \cdot \prod_{s'\in S/\{s\}} \omega(\theta_{s'})\bigg) \end{array} \qquad \begin{array}{r} \forall s \in S, \\ \forall \theta_s \in \Theta_s, \\ \forall \theta'_s \in \Theta_s \end{array} \tag{6}$$

The *ex-ante* weak budget balance constraints are:

$$\sum_{s\in S}\sum_{a\in A} p_{s,a,\theta,r} \geq 0 \qquad \forall \theta \in \Theta, \forall r \in R \tag{7}$$

In what follow we point out the possible objective functions for our model. The maximization of the integrator's *ex-ante* expected utility is:

$$\max \quad \sum_{\theta\in\Theta}\left(\sum_{r\in R}\sum_{s\in S}\sum_{a\in A} p_{s,a,\theta,r} \cdot \overline{\alpha}(a,r)\right) \cdot \prod_{\theta_s\in\theta} \omega(\theta_{s,a}) \tag{8}$$

The maximization of the cumulative search engines' *ex-ante* expected valuations is:

$$\max \quad \sum_{\theta\in\Theta}\left(\sum_{r\in R}\sum_{s\in S}\sum_{a\in A} b(\theta_{s,a}) \cdot f_{s,a,\theta,r} \cdot \bar{\alpha}(a,r)\right) \cdot \prod_{\theta_s\in\theta} \omega(\theta_{s,a}) \tag{9}$$

The maximization of the cumulative search engines' *ex-ante* expected utility is:

$$\max \sum_{\theta \in \Theta} \left(\sum_{r \in R} \sum_{s \in S} \sum_{a \in A} \left(b(\theta_{s,a}) \cdot f_{s,a,\theta,r} - p_{s,a,\theta,r} \right) \cdot \bar{\alpha}(a,r) \right) \cdot \prod_{\theta_s \in \theta} \omega(\theta_{s,a}) \tag{10}$$

The maximization of the search engine s's *ex-ante* expected utility is:

$$\max \sum_{\theta \in \Theta} \left(\sum_{r \in R} \sum_{a \in A} \left(b(\theta_{s,a}) \cdot f_{s,a,\theta,r} - p_{s,a,\theta,r} \right) \cdot \bar{\alpha}(a,r) \right) \cdot \prod_{\theta_s \in \theta} \omega(\theta_{s,a}) \tag{11}$$

It can be easily observed that all the above constraints and objective functions are linear. That is, our formulation is linear mixed integer.

4 Some Examples

We briefly analyze in Section 4.1 the business model currently adopted by AdSense and we show an example of our mechanism in Section 4.2.

4.1 Single Search Engine Case

We consider the business model currently adopted by AdSense where a publisher displays the search engine's sponsored links alongside the search results and, if a link is clicked, the search engine pays a fix ratio of its revenue to the publisher. We study this situation with our framework by introducing a constraint over the payment of the search engine to the publisher (in our case the integrator). Formally, we need to impose that $p_{s,a,\theta,r} = \rho \cdot b(\theta_{s,a})$ with $\rho \in [0,1]$ if $f_{s,a,\theta,r} = 1$. It can be easily shown that no incentive compatible mechanism can be designed in general. Consider the following example.

We have a single search engine s and two possible advertisers. The types related to the first advertiser are $\theta_{s,1} \in \{\theta^1_{s,1}, \theta^2_{s,1}, \theta^3_{s,1}\}$ with $b(\theta^1_{s,1}) = 0.4$, $b(\theta^2_{s,1}) = 0.5$, $b(\theta^3_{s,1}) = 0.6$, and $\alpha(\theta^1_{s,1}) = \alpha(\theta^2_{s,1}) = \alpha(\theta^3_{s,1}) = 0.3$. The types related to the second advertiser are $\theta_{s,2} \in \{\theta^1_{s,2}, \theta^2_{s,2}, \theta^3_{s,2}\}$ with $b(\theta^1_{s,2}) = 0.5$, $b(\theta^2_{s,2}) = 0.6$, $b(\theta^3_{s,2}) = 0.7$, and $\alpha(\theta^1_{s,2}) = \alpha(\theta^2_{s,2}) = \alpha(\theta^3_{s,2}) = 0.2$. The probabilities $\omega(\cdot)$s can be arbitrary. It can be shown that there is no incentive compatible mechanism. Easily, when the true type of search engine s is $(\theta^3_{s,1}, \theta^3_{s,2})$, its optimal strategy is to report $(\theta^1_{s,1}, \theta^1_{s,2})$ independently of the implemented social choice function and independently of the value of ρ. (Practically, our mathematical programming formulation coding the automated mechanism design problem results to be infeasible.) In order to remove this impossibility, we need to remove the constraint on $p_{s,a,\theta,r} = \rho \cdot b(\theta_{s,a})$.

4.2 Analyzing a Case Study

We consider a scenario where an integrator aggregates two domain-specific search engines. The domain of the first search engine is music concerts, while the domain of the second search engine is hotels. A demo of integrator can be found at [7]. We report the user interface of the demo in Fig. 1. We consider a simple example

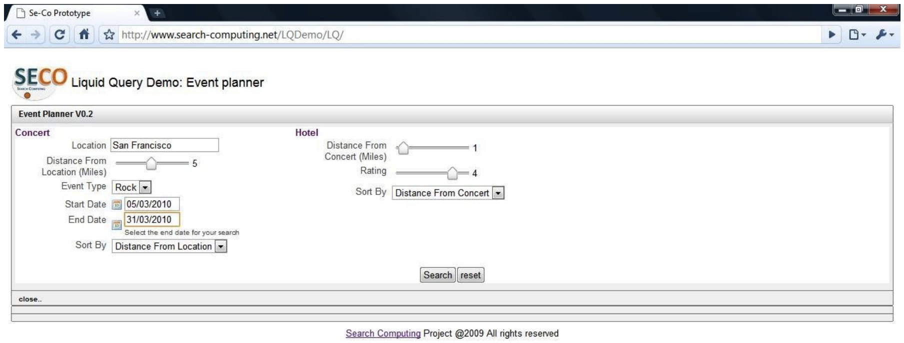

Fig. 1. An example of user interface of an integrator.

and we discuss the functioning of our mechanism (larger scenarios require long time and cannot be solved in exact way in practical applications).

We assume that the user searches for:

– (on the first domain) a concert at Toronto at May 9-15 2010,
– (on the second domain) an hotel at Toronto for the same range of days.

We assume that the first domain-specific search engine returns three sponsored links. We report the Bayesian prior over them in Tab 1. We assume that the second domain-specific search engine returns three sponsored links. We report the Bayesian prior over them in Tab 2. We assume that the number of available slots for sponsored links displayed by the integrator are two.

advertiser	taxi_service				restaurant1				restaurant2	
bid (b)	0.40 €	0.40 €	0.50 €	0.50 €	0.65 €	0.65 €	0.70 €	0.70 €	0.60 €	0.70 €
click probability (α)	0.020	0.030	0.020	0.030	0.040	0.050	0.040	0.050	0.035	0.035
type probability (ω)	0.25	0.25	0.25	0.25	0.30	0.30	0.20	0.20	0.40	0.60

Table 1. Bayesian prior over the sponsored link list returned by the first domain-specific search engine.

advertiser	restaurant1				tourist_office		taxi_service			
bid (b)	0.50 €	0.50 €	0.60 €	0.60 €	0.25 €	0.35 €	0.20 €	0.20 €	0.30 €	0.30 €
click probability (α)	0.030	0.035	0.030	0.035	0.030	0.035	0.020	0.030	0.020	0.030
type probability (ω)	0.30	0.20	0.20	0.30	0.50	0.50	0.25	0.25	0.25	0.25

Table 2. Bayesian prior over the sponsored link list returned by the second domain-specific search engine.

We solved our automated mechanism design problem with Bayes-Nash implementation with the maximization of the integrator's expected utility as objective function. We report the results only for a small number of type profiles. Exactly, we consider the type profiles reported in Tab. 3. The results are reported in Tab. 4, where f_{θ,r_1} denotes the sponsored link displayed in the first position and the search engine it belongs to, f_{θ,r_2} is the same for the second position, p_{θ,r_1} denotes the payment of the search engine whose sponsored link is in the first position once the user clicked on the link, p_{θ,r_2} is the same for the second position. All the other payments are equal to zero. It can be easily observed that the advertiser taxi_service is the one that gives both search engines, if singularly considered, the smallest expected utility. Instead, considering the search engines together, taxi_service is displayed in the second position.

	search engine 1						search engine 2					
	taxi_service		restaurant1		restaurant2		restaurant1		tourist_office		taxi_service	
θ	b	α	b	α	b	α	b	α	b	α	b	α
type profile 1	0.40 €	0.02	0.70 €	0.05	0.60 €	0.035	0.50 €	0.03	0.35 €	0.035	0.20 €	0.03
type profile 2	0.40 €	0.02	0.65 €	0.04	0.60 €	0.035	0.50 €	0.03	0.35 €	0.035	0.30 €	0.02
type profile 3	0.50 €	0.02	0.70 €	0.04	0.70 €	0.035	0.60 €	0.035	0.25 €	0.030	0.20 €	0.03

Table 3. Some type profiles.

θ	f_{θ,r_1}	f_{θ,r_2}	p_{θ,r_1}	p_{θ,r_2}
type profile 1	restaurant1, search engine 1	taxi_service, search engine 1	0.70 €	0.40 €
type profile 2	restaurant1, search engine 1	taxi_service, search engine 1	0.65 €	0.30 €
type profile 3	restaurant1, search engine 1	taxi_service, search engine 1	0.70 €	0.50 €

Table 4. Social choice function and payments.

5 Conclusions and Future Works

The recent advancements in search computing techniques lead to new search paradigms according to which multiple domain-specific search engines are integrated by a special search engine (called integrator). A user can enter a multi-domain query, this query is decomposed by the integrator in a set of single-domain query, each one of them is addressed to a specific-domain search engine. The integrator merges the search results received from each specific-domain search engine. This paradigm allows one to discover a large number of information and to produce very precise search results with respect to the currently available general purpose search engines. In this paper we made a first attempt towards the design of an advertising auction mechanism for this scenario. More precisely, we proposed a business model in which the domain-specific search engines returns, in addition to the search results, a list of sponsored links to the integrator and the integrator merges these lists in a unique list. In order to produce an effective merging, the integrator must be informed about the click probabilities and bids of the advertisers appearing in the lists. We resort to the automated mechanism design framework to design an economic mechanism for the scenario we study. We discuss its desired properties and we report some examples.

The automated mechanism design approach can be used for small problems, but it does not scale for large real-world problem. This pushes for the development of analytical mechanisms or of approximate algorithms. Furthermore, in this paper we have not posed any cooperative constraint over the revenue sharing. In future, our intention is to explore, on the one side, group strategy-proof mechanisms, such as the Moulin's mechanism and its extensions, and, on the other side, two-stage mechanisms to address interdependence valuations.

6 Acknowledgments

This research is part of the "Search Computing" (SeCo) project, funded by the European Research Council (ERC), under the 2008 Call for "IDEAS Advanced Grants", dedicated to frontier research.

References

1. Sakurai, Y., Iwasaki, A., Yokoo, M.: Keyword auction protocol for dynamically adjusting the number of advertisements. In: IAT. (2008) 410–416
2. Narahari, Y., Garg, D., Narayanam, R., Prakash, H.: Game Theoretic Problems in Network Economics and Mechanism Design Solutions. Springer, Berlin, Germany (2009)
3. Google AdWords. (http://adwords.google.com/)
4. Edelman, B., Ostrovsky, M., Schwarz, M.: Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. American Economic Review **97** (2007) 242–259
5. Baeza-Yates, R.A.: Applications of web query mining. In: ECIR. (2005) 7–22
6. White, R.W., Roth, R.A.: Exploratory Search: Beyond the Query-Response Paradigm. Synthesis Lectures on Information Concepts, Retrieval, and Services. Morgan & Claypool Publishers (2009)
7. SECO project. (http://www.search-computing.it/)
8. Google Fusion Tables. (http://tables.googlelabs.com/)
9. Google Squared. (http://www.google.com/squared/)
10. Google AdSense. (http://www.google.com/adsense/)
11. Shoham, Y., Leyton-Brown, K.: Multiagent Systems: Algorithmic, Game Theoretic and Logical Foundations. Cambridge University Press, Cambridge, USA (2008)
12. Sandholm, T.: Automated mechanism design: A new application area for search algorithms. In: CP, Kinsale, County, Springer (2003)
13. Braga, D., Ceri, S., Daniel, F., Martinenghi, D.: Mashing up search services. IEEE Internet Computing **12** (2008) 16–23
14. Bozzon, A., Brambilla, M., Ceri, S., Fraternali, P.: Liquid query: Multi-domain exploratory search on the web. In: WWW. (2010)
15. Dash, R.K., Rogers, A., Jennings, N.R.: A mechanism for multiple goods and interdependent valuations. In: AMEC. (2004) 748–755
16. Mezzetti, C.: Mechanism design with interdependent valuations: Surplus extraction. Discussion Papers in Economics 05/1, Department of Economics, University of Leicester (2005)
17. Krishna, V.: Auction Theory. Academic Press (2006)